

**IMPACTO DE LA ESTRATEGIA DE ACOMPAÑAMIENTO Y SEGUIMIENTO  
ESTUDIANTIL ASES EN EL RENDIMIENTO ACADÉMICO DE LOS  
ESTUDIANTES DE SER PILO PAGA EN LA UNIVERSIDAD DEL VALLE: UNA  
ESTIMACIÓN POR PROPENSITY SCORE MATCHING**

**GUIDO MAURICIO ALVARADO PLATA**

**UNIVERSIDAD DEL VALLE  
FACULTAD DE CIENCIAS SOCIALES Y ECONÓMICAS  
PROGRAMA ACADÉMICO DE ECONOMÍA  
SANTIAGO DE CALI  
2019**

Impacto de la estrategia de acompañamiento y seguimiento estudiantil ASES en el  
rendimiento académico de los estudiantes de Ser Pilo Paga en la Universidad del Valle: una  
estimación por *Propensity Score Matching*

Guido Mauricio Alvarado Plata

Trabajo de grado para optar al título de Economista

Director  
Jaime Humberto Escobar Martínez

Universidad del Valle  
Facultad de Ciencias Sociales y Económicas  
Programa Académico de Economía  
Santiago de Cali

2019

## TABLA DE CONTENIDO

	Pág.
RESUMEN	
1. INTRODUCCIÓN	2
2. REVISIÓN DE LITERATURA	4
3. MARCO TEÓRICO	7
4. PROGRAMA ASES	14
5. MÉTODOS Y PROCEDIMIENTOS	16
5.1 Metodología de <i>Propensity Score Matching</i>	17
5.2 <i>Propensity Score Matching</i> con datos agrupados	21
5.3 Estructura de los datos	23
5.3.1 Indicador de Esfuerzo Académico	25
5.4 Modelo	27
5.4.1 Estimación de la probabilidad de participación	28
5.4.2 Resultados modelos <i>Propensity Score Matching</i>	30
5.4.3 Prueba de sensibilidad de Rosenbaum	32
6. CONCLUSIONES	35
7. REFERENCIAS	37

## RESUMEN

La investigación busca encontrar el impacto del programa de Acompañamiento y Seguimiento Estudiantil (ASES) sobre el rendimiento académico de los estudiantes Ser Pilo Paga-II de la Universidad del Valle durante el ciclo básico utilizando la metodología de *Propensity Score Matching*. Este ejercicio guarda relevancia por su interés en ofrecer una luz para orientar política universitaria y analizar el papel del rendimiento académico en la deserción estudiantil que aqueja las universidades públicas de Colombia. Se encontró que, a partir del primer año de intervención, el rendimiento académico de los estudiantes intervenidos por ASES fue superior respecto al rendimiento académico de los estudiantes no intervenidos.

**Palabras claves:** Economía de la educación, Deserción universitaria, ASES, Rendimiento académico, Propensity Score Matching.

**JEL:** I20, I21, I23.

## **1. INTRODUCCIÓN**

La deserción estudiantil es un fenómeno que aqueja a la educación colombiana en su conjunto, tanto a nivel de secundaria como a nivel superior, y que altera los procesos de formación de capital humano del país comprometiendo requisitos básicos para el crecimiento y desarrollo socioeconómico. El alcance de este problema no es meramente institucional, sino que este fenómeno se refleja además en aspectos socioeconómicos, académicos e individuales pues los principales afectados son los estudiantes que ven truncados sus proyectos de vida (Ministerio de Educación Nacional, 2014)

Según Guzmán et al. (2009), las políticas de educación que se han implementado en el país van dirigidas a aumentar la cobertura de la educación superior, debido a que la demanda de educación superior es mucho mayor que la oferta. No obstante, a la hora de ver los resultados de esta política se encuentra que hasta que no se controle el fenómeno de deserción, dicha política no tendrá el impacto esperado.

En este sentido, diversos autores han abordado la problemática buscando cuáles serían los determinantes de la deserción universitaria, encontrando estudios como los de Girón & González (2005) y Vásquez, et al. (2003), en los que en general, se plantea que este fenómeno es multidimensional siendo menester realizar estudios para casos específicos pues los determinantes pueden ir variando de acuerdo con el contexto universitario.

La Universidad del Valle no es ajena a este problema y se ha puesto en la tarea, a lo largo de los años, de encontrar cuáles podrían ser los determinantes de la deserción de la institución y una vez identificados, preguntarse qué se podría implementar para disminuir la magnitud de este fenómeno. Escobar, et al. (2006), analizan los factores determinantes de la deserción en la Universidad del Valle encontrando que el rendimiento académico es clave para explicar la deserción. En esta vía, el Ministerio de Educación Nacional constata este hallazgo empírico al plantear que el principal factor determinante de la deserción en Colombia se debe a cuestiones académicas (Guzmán et al., 2009)

Por lo anterior, en la Universidad se vio la necesidad de preguntarse: ¿qué se podría implementar para darle un giro al fracaso académico? En este sentido, surgió la idea de un mecanismo en el que fuera posible una intervención para ayudar a que este fenómeno no siga propagándose. Dicha intervención se realiza por medio de la estrategia ASES (Acompañamiento y Seguimiento Estudiantil) para los estudiantes de nuevo ingreso que sean beneficiarios de Ser Pilo Paga o admitidos por condición de excepción, con el fin de proyectar la estrategia a todas las cohortes venideras para que la deserción estudiantil asociada a un bajo rendimiento académico disminuya considerablemente y el éxito académico sea más probable. Sin embargo, hasta el momento no se han realizado evaluaciones de los resultados del programa, siendo necesario estimar si el programa ha tenido incidencia en el rendimiento académico de los estudiantes y, por consecuencia, ha logrado influir en la reducción del fracaso académico.

En el presente trabajo se da espacio para responder a la pregunta de investigación: ¿cuál es el impacto de la estrategia ASES en el desempeño académico en el ciclo básico de los estudiantes Ser Pilo Paga-II admitidos en el primer semestre de 2016 a la Universidad del Valle? Para alcanzar la respuesta, los escalones necesarios son analizar cuáles son los determinantes de la deserción estudiantil en la Universidad del Valle, examinar si la estrategia ASES permite el acercamiento al éxito académico, esclarecer la importancia del capital humano en las mejoras de desarrollo del país e identificar la influencia que tienen las variables individuales sobre la deserción para así lograr llegar al objetivo general de la investigación que consiste en identificar el impacto de la estrategia ASES en el desempeño académico de la población intervenida.

Para lograr resolver el interrogante se utiliza la metodología *Propensity Score Matching* de la evaluación de impacto, siendo este un análisis ex -post de la estrategia implementada desde el primer semestre del 2015.

Este ejercicio guarda relevancia por su interés en ofrecer una luz para orientar política universitaria y mejorar el rendimiento de los estudiantes recién admitidos a la Universidad del Valle, teniendo en cuenta que, de acuerdo con los antecedentes de educación, la deserción estudiantil se produce principalmente en el ciclo básico de estudio y resulta ser

el desempeño de los estudiantes el determinante que en nuestro contexto se ha identificado como el de mayor peso. En una fase inicial se estaría evitando que se desaprovechen los cupos ofrecidos y que un mayor número de estudiantes logren el éxito académico teniendo una adecuada tasa de retención estudiantil (Lopera, 2008); por otro lado, se estarían evitando huecos financieros y desperdicio de recursos públicos (Tinto, 1989) y una última dimensión en la que se obtendrían resultados positivos es en el desarrollo personal del individuo, se estaría fomentando la inversión en capital humano que se reflejará en mejores indicadores de desarrollo para el país (De Escobar, et al., 2005). Complementando a esta última dimensión hay que tener en cuenta que aquellos estudiantes Ser Pilo Paga estarían evitando que su beca se vuelva un crédito al no cumplir con los requisitos por parte del Ministerio de Educación de culminar su programa de estudio en los plazos establecidos.

El ejercicio encuentra que el programa ASES es una estrategia efectiva que ha logrado un impacto en el rendimiento académico de los estudiantes intervenidos; de esta forma, se lograría contrarrestar el principal determinante de la deserción en la Universidad del Valle mejorando los indicadores de deserción en el ciclo básico de estudio.

Este reporte de investigación está organizado de la siguiente forma: la presente sección introductoria, luego se expone la revisión de literatura, acto seguido se plantea el marco teórico, posteriormente se desarrollan los métodos y procedimientos y, por último, se plantean recomendaciones y conclusiones.

## **2. REVISIÓN DE LITERATURA**

En la primera clase de estudios analizados, se encontraron autores que hacen un análisis general de los determinantes de la deserción escolar para universidades del país. Por un lado, Vásquez, et al. (2003) exponen un modelo de duración que estima los principales factores de deserción y calculan el riesgo de desertar al transcurrir los semestres para la Universidad de Antioquia; el fenómeno de esta Universidad se concentra en la llamada deserción temprana (de 1-4 semestre) y es inferior la deserción precoz y tardía. Otro hallazgo relevante de su estudio es que la orientación vocacional previa al ingreso a la Universidad es fundamental para disminuir el riesgo de deserción y que el hecho de ser

hombre aumenta el riesgo de desertar.

Por otro lado, Lopera (2008) hace el análisis por medio de un modelo de duración de riesgo proporcional en tiempo discreto para la Universidad del Rosario, encontrando igualmente que el riesgo de desertar es mayor en el núcleo básico apuntando al rendimiento académico como el determinante de dicha deserción; otros de los factores claves para su estudio son el sexo, estar vinculado al mercado laboral a la par de estar estudiando y provenir de otra región diferente a la que está localizada la universidad.

Ahora bien, concierne a los antecedentes que se han enfocado en el rendimiento académico como el determinante de su investigación tenemos a Porto & Di Gresia (2004) que utilizan una regresión múltiple para explicar los determinantes del bajo rendimiento académico en la Facultad de Ciencias Económicas de la Universidad de la Plata, encuentran que guardan un efecto positivo las variables de nivel educativo de los padres, desempeño en la secundaria y el hecho de ser mujer; por su parte, guardan un efecto negativo en el rendimiento las variables tales como el hecho que el estudiante trabaje simultáneamente en el periodo académico y la edad con que ingresó a la Universidad (entre más joven ingrese, mejor rendimiento).

En otro estudio, Girón & González (2005) también parten del hecho de que el rendimiento académico determina la deserción estudiantil y por tanto, buscan analizar las variables que tienen influencia en el fenómeno de la deserción para la Universidad Pontificia Javeriana en el que se encuentra nuevamente que el sexo es una variable que influye en el rendimiento académico, el número de créditos matriculados es otra de las variables que resultan significativas para este estudio al igual que el rendimiento previo de pruebas en matemáticas y lenguaje.

Bajo esta lógica, Escobar et al. (2006) encuentran que la Universidad del Valle no es ajena al problema de la deserción estudiantil e identificaron que el principal determinante de la deserción en la institución está ligado al rendimiento académico. Los autores concuerdan con los estudios anteriormente expuestos al plantear que tanto a nivel nacional como internacional el fenómeno de la deserción es mayormente probable para personas de sexo masculino y que un adecuado mecanismo de incentivos como acompañamiento y acceso a programas de bienestar universitario podría ser clave para enfrentar el problema. En este



mismo sentido, el Ministerio de Educación Nacional a través del estudio realizado por Guzmán et al. (2009) constata este hallazgo empírico de los autores al plantear que “Los recientes resultados del seguimiento de la deserción estudiantil en educación superior demuestran que el principal factor determinante del abandono de estudios en Colombia se sitúa en la dimensión académica”

Ahora bien, respecto a los antecedentes que dieron paso a la creación del programa ASES se encuentra el trabajo de Escobar et al. (2008) en el que se plantean diez estrategias de política universitaria para hacerle frente al problema de deserción, sobrepermanencia y no graduación en la Universidad del Valle. Los lineamientos destacan la necesidad de un programa de acompañamiento socioeducativo coordinado por la misma Universidad en el que se logre alcanzar la sintonía del estudiante con el ambiente universitario, que motiven al estudiante a permanecer en la Universidad y además, se cuente con un programa de consejería estudiantil que ayude a orientar las dificultades por las que puedan pasar los estudiantes facilitando técnicas de estudio, bibliografía alternativa, aplicaciones de los contenidos, servicio psicológico, entre otras.

El último factor de antecedentes a abordar gira en torno a las políticas y estrategias realizadas en el sector educativo; aquí se destacan estudios como el de Angrist et al. (2006) que realizan una evaluación de impacto de una política de cobertura de educación secundaria en Bogotá llamada PACES (Programa de Ampliación de Cobertura de la Educación Secundaria) que se implementó en el año 1991; la estrategia consistía en hacer entrega de unos *vouchers* para que los discentes estudiaran en un colegio privado y posteriormente medir el progreso académico de los mismos. A partir de ello encontraron que el porcentaje de graduación para bachilleres tratados aumentó entre 15% y 20% gracias a la estrategia.

Otro estudio propuesto por Ramírez (2013) mide el impacto de la estrategia ICARO (Incentivos Condicionados al Acceso y a la Retención Oportuna, en la educación secundaria y media) en la deserción escolar de tres ciudades del país (Buenaventura, Montería y Cúcuta) para el año 2010. El programa permite que el beneficiario escoja entre un número de colegios privados seleccionados previamente con la motivación de que sea posible el acceso a educación de calidad e infraestructura óptima a poblaciones que de

acuerdo con su nivel de ingresos normalmente no podrían acceder, por tanto, se tendría la creencia de que la educación privada en el país brinda mejores incentivos para la retención estudiantil. El resultado para esta evaluación es que la deserción intra anual disminuyó entre 1,95 y 2,64 puntos alcanzando los objetivos propuestos por el autor.

Por último, se encuentra la evaluación de impacto de una estrategia destinada también a estudiantes de escasos recursos en el que a diferencia del programa anterior, se les ofrece la oportunidad de hacer parte de una institución pública pero que será manejada por directivas privadas con experiencia en educación por un lapso de 15 años, a esta estrategia se le llamó CEC (Colegios en Concesión); se encuentra que la deserción de los estudiantes acogidos por la estrategia es de 1,7% menos que para otros colegios públicos, resultado obtenido gracias a que las concesiones planificaron su actuar en pro de ofrecer apoyo psicológico y visitas domiciliarias de profesionales como trabajadores sociales y psicólogos cuando identificaban señales de alerta de deserción. (Barrera, 2007)

Tras analizar los antecedentes que hacen un acercamiento a la problemática de la deserción escolar es posible relacionar los hallazgos de los diversos autores y esclarecer el panorama respecto al fenómeno de la deserción. Por tanto, según el análisis de los autores, resultaría relevante para la investigación identificar los incentivos adecuados para dinamizar el logro académico; además se tiene en cuenta que los estudios de evaluación de impacto ayudan a analizar si los esfuerzos por llevar a cabo diferentes tipos de política han sido significativos en sus objetivos aislando diferentes factores que pueden influir y concentrando el análisis en el efecto de la variable que recoge el tratamiento propuesto.

### **3. MARCO TEÓRICO**

La deserción es un concepto difícil de describir al poseer aspectos multidimensionales, por lo que podría considerarse como una cuestión de perspectiva; puede ser vista como un acto voluntario o involuntario (Tinto, 1989). El trabajo de Himmel (2002) recopila los diferentes enfoques desde los cuales se ha estudiado el fenómeno de la deserción tales como: psicológicos, sociológicos, organizacionales, económicos y de interacciones.

Desde el enfoque psicológico, se explica la deserción a partir de las diferencias en la personalidad de los estudiantes que alcanzan a culminar sus estudios y los desertores, es decir, se concentra en variables que giran en torno a las características propias del individuo (Himmel, 2002). Este tipo de análisis partió del modelo propuesto por Ajzen & Fishbein (1975) en donde se plantea que el comportamiento de los individuos está influenciado por creencias normativas subjetivas y por lo que se cree será consecuencia de su actuar, es decir, la deserción es producto de un cambio o debilitamiento de las intenciones que se tenían al iniciar los estudios.

Attinasi (1986) complementa el anterior enfoque introduciendo la idea de que otro factor de incidencia en la deserción es la percepción por parte de los estudiantes de su experiencia en el ámbito universitario.

Eccles et al. (1983), por su parte, plantean la importancia del *background* familiar en este tipo de análisis, en otras palabras, la influencia de los padres en los hábitos de sus hijos. En este sentido, el apoyo por parte de los padres es una fuerte influencia en los logros de los hijos. Además, expone que el rendimiento previo del estudiante influencia en el rendimiento futuro afectando el autoconcepto del individuo, es decir, cambia la percepción de dificultad de los estudios, de sus metas y sus expectativas de éxito.

Tiempo después, Ethington (1990) siguió complementando el enfoque psicológico y logra compactar una teoría más general basado principalmente en “las conductas del logro”, las cuales abarcan atributos de los individuos tales como sus elecciones, desempeño y la perseverancia. Plantea que tanto el apoyo brindado por parte de la familia como el rendimiento académico previo del estudiante determina la percepción de dificultad de los estudios, esto a su vez, puede provocar alteraciones en el nivel de las aspiraciones propias del individuo, lo que desembocaría en posibilidades de desertar de los estudios universitarios.

Desde el enfoque sociológico, la teoría de deserción de Spady (1970) plantea que la deserción debería intentar explicarse desde un enfoque interdisciplinario más riguroso que tenga en cuenta las interacciones entre el estudiante como individuo y el ambiente institucional ya que dicha interacción da paso a que los atributos del estudiante estén expuestos a influencias y expectativas que deben tenerse en cuenta en el análisis. Dada la

variedad de opciones para analizar la problemática, la teoría contempla la deserción como el acto en que cualquier estudiante abandone una universidad en la que se encontraba matriculado.

Además, se incorpora el rendimiento académico del estudiante al modelo, encontrándose inseparablemente relacionada con la deserción, debido a que, por políticas institucionales, una gran parte de los abandonos universitarios se debe a los bajos rendimientos.

Respecto a los supuestos planteados, las recompensas tanto académicas como sociales juegan un papel clave ya que la satisfacción de los estudiantes con la experiencia universitaria dependerá de ellas y, en este sentido, se definirá la integración al sistema institucional.

Spady (1970) utiliza el modelo de suicidio de Durkheim (1951) para hacer un paralelo al contexto de la deserción entre la denominada integración social y el éxito académico. Plantea que el éxito puede ser visto desde dos factores; el primero es llamado congruencia normativa refiriéndose a la compatibilidad entre actitudes, intereses y disposiciones de personalidad y los atributos e influencias del entorno, la segunda es llamada apoyo de amistad, es decir, el tejer relaciones cercanas con los demás individuos en el sistema.

En este enfoque sociológico, el *background* familiar proporciona la base para el sistema ya que influencia tanto el potencial de rendimiento como la congruencia normativa, es decir, las consecuencias de la interacción de los atributos propios del individuo y su entorno universitario. En este sentido, se hace hincapié en que el desempeño del estudiante juega un papel complejo en el análisis ya que puede llegar a anular el patrón teórico del modelo, en otras palabras, el fracaso por parte del estudiante hará que por reglas institucionales se vea obligado a retirarse, aunque su nivel de satisfacción o compromiso con la institución sea alto.

Ahora bien, pasando al enfoque organizacional se encuentra un punto de vista institucional, la teoría de Braxton, et al. (2000) plantea que la deserción es explicada por la conformidad del estudiante respecto a las características de la institución de educación superior, es decir, su percepción frente a la calidad de la docencia y su experiencia en el aula de clases. En este mismo sentido, Tillman (2002) plantea que dentro de los factores de influencia se encuentran el acceso que tengan los estudiantes a laboratorios, la

disponibilidad de recursos bibliográficos y demás facilidades que ofrezca la institución educativa.

Lasibille & Navarro (2004) se concentran en explorar las herramientas necesarias para una planificación del sistema educativo ya que esto permitiría concentrarse en factores de cantidad de educación ofrecida y calidad de la misma, además de responder a las necesidades del crecimiento económico y su respectiva demanda social en pro de aumentar la equidad en la sociedad gracias a una planificación adecuada.

Por su parte, Tinto (1989) también hace un aporte a este enfoque de análisis planteando que, desde el punto de vista organizacional, todo estudiante que abandone sus estudios podrá ser catalogado como desertor ya que se crea un lugar vacante en el conjunto estudiantil que pudo haber sido aprovechado por otro individuo que sí persistiera en los estudios hasta alcanzar su título.

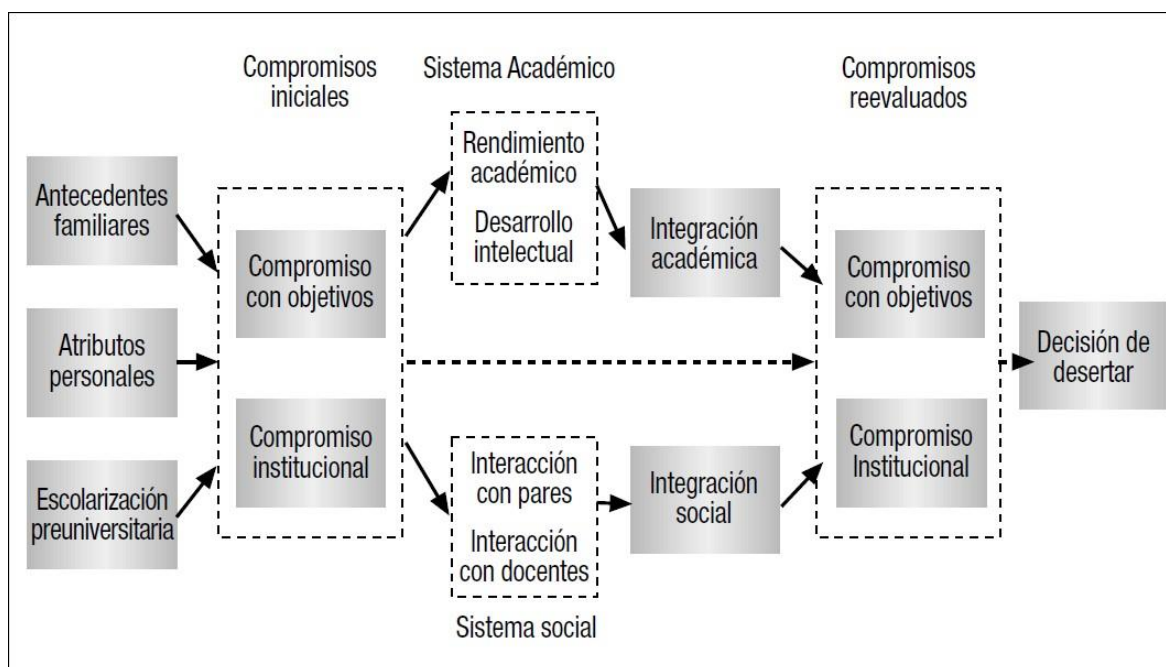
En el enfoque económico, Tinto (1975) desarrolla el análisis costo-beneficio partiendo de la racionalidad de los individuos, el estudiante realiza el cálculo de si el beneficio alcanzado al final de su formación será mayor que el costo de oportunidad asumido al no estar invirtiendo su tiempo en trabajo u otras actividades en el corto plazo.

Además de ello, el enfoque contempla en su teoría la incapacidad del estudiante por asumir los costos asociados a los estudios (Himmel, 2002). Por tanto, John, et al. (1999) atendiendo a este punto de la incapacidad para solventar los costos del estudio, plantean la influencia de los subsidios estudiantiles en el fenómeno de la deserción. Si los beneficios estudiantiles son bien focalizados, el impacto será efectivo y positivo en la retención de estudiantes.

Una de las teorías más influyentes en los análisis de deserción según Himmel (2002) es la realizada por Tinto (1975) ya que recoge diferentes aspectos de interacción entre el individuo y la institución; de este modo, podría catalogarse dentro del enfoque de las interacciones. Se considera que a medida que el estudiante se sumerge en la vida universitaria, la adaptación a la institución de educación superior estará determinada por características previas al ingreso como los antecedentes familiares, los atributos personales

y cómo fue su escolarización previa influyendo sobre el compromiso del estudiante y sus metas. Si se logra una integración, no solo a nivel social sino también académica (rendimiento y desarrollo de habilidades) será un aspecto fundamental del modelo ya que gracias a este proceso se logra la reevaluación de los compromisos en el periodo inicial y de acuerdo con ello se tomará la decisión de continuar los estudios o desertar. Concluye que habrá una menor probabilidad de desertar si el estudiante consolida un compromiso con la culminación de sus estudios y con la institución en la que se formará mejorando su integración social y rendimiento académico. Lo anterior se puede visualizar en el Gráfico 1.

**Gráfico 1. Modelo de Tinto (1975)**



**Fuente:** Modelos de análisis de la deserción estudiantil en la educación superior, E. Himmel (2002)

Por su parte, Bean (1985) amplía el modelo de Tinto (1975) haciendo un paralelo con la productividad laboral; la satisfacción en los estudios es similar a la satisfacción con el trabajo. Combina el enfoque psicológico en el que se planteaba que las creencias del individuo influyen sobre las intenciones conductuales, siendo estas últimas predictores de la deserción y expone que dichas creencias son afectadas a su vez por características institucionales que se planteaban en el enfoque organizacional como la calidad de la educación recibida y los tipos de docentes. Entre los aportes de Bean (1985) se incorpora al análisis los factores ambientales que explican la decisión de desertar al ser el efecto de relaciones sociales externas en la adaptación a la vida universitaria. Por tanto, para el autor, factores no cognitivos (características propias del individuo), factores ambientales (financiamiento, oportunidades de transferencia) y factores organizacionales (desempeño académico, integración académica), tendrán significancia a la hora de explicar el fenómeno de la deserción.

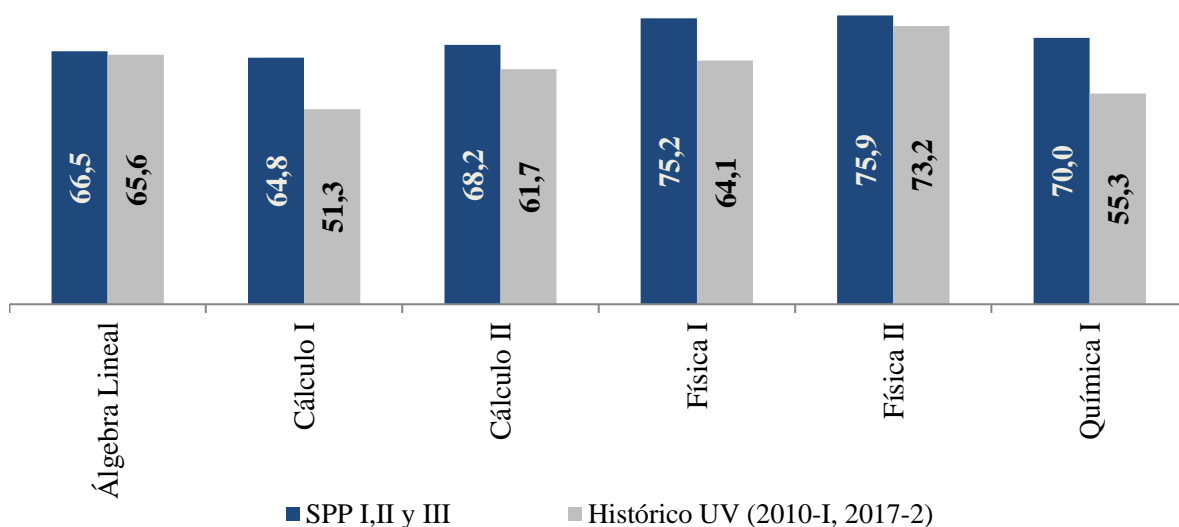
De este modo, John et al. (1999) sintetizan los aportes de los autores anteriormente mencionados y proponen una teoría más compacta en la que se plantean tres etapas; la primera etapa contempla la disposición del estudiante por seguir con sus estudios según sus antecedentes de habilidades académicas y factores socioeconómicos. La segunda etapa consiste en el análisis costo-beneficio que realiza el alumno para decidir si culminar los estudios es la decisión más racional. La tercera y última etapa se ubica en la situación en que el estudiante ya inició sus estudios y de acuerdo con sus experiencias sociales y académicas, sus aspiraciones pueden sufrir alteraciones.

Por último, Weidman (1989) aporta al enfoque de interacciones al plantear que la socialización de los estudiantes en su programa académico depende de factores tanto psicológicos, sociológicos e institucionales, por tanto, asume que los estudiantes inician su proceso de formación con unas variables dadas tales como sus aspiraciones, valores, nivel socioeconómico, aptitudes, presiones de la sociedad, interés de estudiar, entre otras.

A manera de síntesis, de acuerdo con el análisis teórico proporcionado por los autores sobre la deserción estudiantil, el fenómeno de deserción es complejo ya que se puede definir por un conjunto de interacciones humanas: dimensiones sociológicas, económicas, religiosas, históricas, biológicas, etc., que están en constante interacción (Morín, 1999). Esta postura la comparte Himmel (2002) al plantear que independientemente de la decisión de qué enfoque utilizar para abordar el problema de investigación, este desembocará igualmente en apreciar aspectos como las características del individuo, de lo académico, de las instituciones y lo socioeconómico.

De acuerdo con la teoría expuesta, la hipótesis que se maneja en la investigación es que la estrategia ASES tiene un impacto positivo en el rendimiento académico de los estudiantes intervenidos ya que, desde un análisis conceptual, el trasfondo del programa responde a lo planteado por Spady (1970) al atender la congruencia normativa y el apoyo de amistad. Por otro lado, desde un análisis descriptivo también se sospecha tal efecto positivo al aumentar el porcentaje de aprobación de materias críticas del ciclo básico (Gráfico 2) y la reducción del porcentaje de desertores tempranos para la población bajo estudio (Tabla 1).

**Gráfico 2. Niveles de aprobación (%) en asignaturas críticas (menos de 80% de aprobación) en estudiantes Ser Pilo Paga (I, II y III) Vs. históricos Univalle (2010-1 a 2017-2)**



**Fuente:** Oficina de ASES con base en información de registro académico – Elaboración propia



**Tabla 1. Comparativo entre deserción institucional y poblaciones atendidas por ASES**

Población	Semestre					
	1	2	3	4	5	6
<b>Univalle (2009-2016)</b>	11,8%	19,5%	25,3%	29,5%	32,8%	35,6%
<b>SPP1</b>	1,7%	8,7%	10,4%	13,9%	16,5%	18,3%
<b>SPP2</b>	3,2%	6,4%	8,7%	11,1%		
<b>SPP3</b>	6,1%	8,5%				

**Fuente:** Oficina de ASES con base en información de registro académico – Elaboración propia

#### **4. PROGRAMA ASES**

El programa ASES es una estrategia propuesta por la Universidad del Valle para promover la adaptación de los estudiantes que ingresan bajo Condiciones de Excepción y Ser Pilo Paga a la vida universitaria. Esto implica un seguimiento a componentes sociales y académicos de los estudiantes, buscando potenciar el desarrollo de habilidades, conocimientos, actitudes y hábitos acordes con lo requerido en los programas académicos de la Universidad.

El programa cuenta con un componente Socioeducativo que está encargado de liderar el acompañamiento a los estudiantes y monitorear posibles alertas de riesgo de deserción en dimensiones académicas, económicas, familiares, individuales o de vida universitaria. Para ello, se contratan monitores socioeducativos que se encargan de transmitir experiencias y acompañar entre cinco a diez estudiantes de su mismo programa académico y registrar todas aquellas circunstancias por las que atraviesa el estudiante.

Existe otro componente llamado Gestión de la Información; en este, se organiza la información obtenida por el componente socioeducativo y a través de herramientas virtuales se visualiza en tiempo real la información emitiendo señales de alerta cuando la situación lo amerite, ya sea una ruta de atención de acompañamiento académico, atención profesional o de bienestar, salud y deporte.

La estrategia también realiza una evaluación permanente del proceso que viven los estudiantes por medio del componente de Investigaciones. Finalmente, el componente académico coordina la realización de monitorías académicas, pruebas diagnósticas y junto al proyecto GRACA (Grupos de Apoyo a la Cultura Académica) se brinda apoyo a estudiantes con temas referentes a la lectura y escritura.

Barbosa, Castillo & Vásquez (2018) plantean que las principales acciones que realiza la Estrategia ASES son:

- **Acompañamiento entre pares:** se asigna un monitor de la misma carrera a cada estudiante que ingresa a ASES; su función es orientar al estudiante sobre los diferentes retos que enfrentará en la Universidad y su programa académico. A su vez, el monitor está encargado de recopilar información sobre situaciones de riesgo que enfrenta el estudiante y las consigna en una plataforma digital.
- **Acompañamientos grupales:** se busca crear una red de apoyo que le permita al estudiante sentirse acogido y compartir sus experiencias e inquietudes tanto con sus compañeros como con los monitores.
- **Seguimiento al desempeño académico:** permite detectar situaciones de riesgo académico a partir de las cuales se activan diversas acciones de apoyo que eviten la ocurrencia de situaciones de pérdida de asignaturas, ocurrencia de bajos rendimientos y finalmente fracaso académico.
- **Monitorías académicas individuales y grupales:** realizadas en las asignaturas que históricamente han tenido los más altos índices de reprobación. Estas no solo buscan resolver vacíos conceptuales sino también promover un papel activo de los mismos estudiantes en su aprendizaje, el desarrollo de hábitos y diversas habilidades –entre ellas las comunicativas- que permitan al estudiante hacerse cargo de su propio proceso formativo.
- **Orientaciones individuales:** cuando las situaciones trascienden las posibilidades de resolución entre el monitor y el estudiante actúan directamente los profesionales. En ellas se orienta al estudiante sobre las diferentes alternativas a las que puede recurrir en cada situación.
- **Georreferenciación:** permite conocer la ubicación de la vivienda del estudiante, la distancia con la Universidad, los niveles de riesgo asociados al contexto, como factores que pueden estar asociados con la posibilidad de deserción. Cuando se requiere, se ayuda y orienta a los estudiantes a encontrar una vivienda que favorezca dichas variables.

## 5. MÉTODOS Y PROCEDIMIENTOS

La investigación busca encontrar el impacto del programa ASES sobre el rendimiento académico de los estudiantes Ser Pilo Paga-II de la Universidad del Valle. Para este tipo de evaluación, se requiere implementar una metodología que permita establecer una relación causal entre el programa ASES y el impacto en el rendimiento académico por medio del efecto del tratamiento que se obtiene de la diferencia entre la variable de resultado del individuo que participa en el programa y la variable de resultado del individuo en caso de que no existiera el programa.

Es necesario, por tanto, identificar dos tipos de poblaciones: una población intervenida (aquellos bajo el tratamiento o programa) y una población de control (aquellos que no hacen parte del programa, pero pudieron haber sido elegibles), con el fin de utilizar el resultado de los no participantes (pero elegibles) como aproximación de los resultados que habrían obtenido los participantes en caso de no existir el programa ya que no es posible observar los dos resultados para un mismo individuo, a esto se le llama el contrafactual (Bernal & Peña, 2011).

La presente investigación está catalogada bajo la gama de estudios no experimentales de la evaluación de impacto; este tipo de trabajos se caracterizan porque los datos disponibles no provienen de una asignación aleatoria al programa, es por ello que es necesario plantear una serie de supuestos que ayuden a solucionar el sesgo de selección que por construcción tienen los estudios no experimentales y que constituye la principal limitación de la metodología. Los estudios no experimentales pueden generar estimaciones inexactas del impacto del programa, ya sea sobreestimándolo o subestimándolo debido a que los participantes y no participantes generalmente son diferentes, pueden existir diferencias tanto observables como no observables entre los dos grupos (Bernal & Peña, 2011).

Como se planteó anteriormente, para calcular el efecto del programa en los individuos tratados se analiza la diferencia entre la variable de resultado del individuo que participó en el programa y la variable de resultado del individuo en caso de que no existiera el programa (contrafactual). Lo anteriormente expuesto es abarcado en el modelo Roy-Rubin y este se expresa como:

$$E(Y_i(1)|D_i=1) - E(Y_i(0)|D_i=0)$$

Donde  $Y_i(0)$  y  $Y_i(1)$  son los resultados potenciales para cada grupo, 1 para tratamientos y 0 para controles. El contrafactual es tal que  $E(Y_i(0)|D_i=1)$ .

Se involucra en la ecuación el contrafactual como un 1 y se obtiene:

$$=E(Y_i(1)|D_i=1) - E(Y_i(0)|D_i=1) + E(Y_i(0)|D_i=1) - E(Y_i(0)|D_i=0) \quad (1)$$

La expresión denotada en (1) muestra en sus dos primeros términos el *Average Treatment on the Treated* (ATT), es decir, el efecto que tiene el programa sobre los resultados de los individuos tratados menos los resultados que habrían obtenido dichos individuos si no hubieran sido parte del programa. Los dos últimos términos expresan el sesgo de selección, es decir, la diferencia entre la variable de resultado que se habría obtenido si el individuo participara en el programa y la variable de resultado de los individuos que no hicieron parte del programa.

### 5.1 Metodología *Propensity Score Matching*

La metodología no experimental llamada *Propensity Score Matching* (PSM) consiste en encontrar un “clon” de los individuos que están bajo tratamiento en la población de control. A lo largo del tiempo, el estudio del método de emparejamiento experimentó grandes avances; en 1983 tuvo lugar un importante suceso que fue la introducción de los *Propensity Score* definido por Rosenbaum & Rubin (1983) como la probabilidad de recibir el tratamiento dado un conjunto de covariables observables.

Los *Propensity Score* según Luellen, et al. (2005) se crean mediante una regresión logística y muestran la probabilidad estimada de que un individuo participe en el programa dado un conjunto de covariables. En este mismo sentido, Joffe & Rosenbaum (1999) plantean que son la probabilidad condicional de exposición a un tratamiento de acuerdo con unas variables observables dadas. A diferencia de la estructura anteriormente expuesta en la sección 5, los *Propensity Score* facilitaron el emparejamiento de los grupos de tratamiento y control debido a que requieren coincidencias cercanas por medio de un escalar y no necesitan de coincidencias exactas en el conjunto de todas las covariables que se estén utilizando, que pueden llegar a ser dimensiones muy grandes.

El estimador del PSM se expresa de la siguiente forma:

$$P(X) = P(D=1|X) \quad (2)$$

La expresión (2) muestra que la probabilidad de participar en el programa estará determinada por las características observables, en este sentido, Bernal & Peña (2011) plantean que la probabilidad de participación  $P(X)$  resume todas las variables observables de un individuo. Stuart (2010) planteó que la agrupación de individuos con *Propensity Score* similares replica un mini-experimento aleatorizado, al menos con respecto a las covariables que son observadas.

Retomando la ecuación (2), Rosenbaum & Rubin (1983) plantean que para lograr deshacerse del sesgo de selección que por construcción hace parte de esta metodología de PSM, se recurre a tener en cuenta dos supuestos claves: la independencia condicional y el soporte común.

El supuesto de independencia condicional se denota como:

$$Y(0), Y(1) \perp D|P(X), \forall X \quad (3)$$

Esta condición establece que la selección en el programa está basada solamente en variables que son observables, se excluyen del análisis aquellas variables no observables que puedan influenciar tanto en la participación del programa de los individuos como en la variable de resultado potencial ( $Y_i$ ).

En otras palabras, al condicionar que la participación en el programa estará determinada únicamente por variables que son observables y que serán incluidas en el modelo, garantizará que las diferencias en el promedio de las variables de resultado de los dos grupos se deben exclusivamente al efecto que tiene el programa sobre la variable de resultado de los individuos, eliminando de esta forma, el sesgo de selección.

Adicional al supuesto de independencia condicional, la condición de soporte común también hace parte de los supuestos claves del PSM; este hace referencia a que los individuos deben tener probabilidades positivas de ser elegibles tanto para participar en el programa como para ser parte del grupo de control (Heckman, LaLonde & Smith, 1999).

Por tanto, el restringir la muestra al soporte común es el mecanismo para garantizar que se están comparando grupos que efectivamente son comparables, individuos del grupo de control que tengan probabilidades de participar en el programa similares a las de los individuos del grupo de tratamiento. El cumplimiento de estos supuestos garantiza según la metodología que nuestros resultados serán robustos.

La condición de soporte común se expresa de forma tal que:

$$0 < P(D=1|X) < 1 \quad (4)$$

Al cumplirse los supuestos del PSM, el estimador ATT se expresa como:

$$ATT_{PSM} = E_{P(X)|D=1} \{E[Y(1)|D=1, P(X)] - E[Y(0)|D=0, P(X)]\} \quad (5)$$

Donde  $E_{P(X)|D=1}$  es el valor esperado respecto a la probabilidad de participación  $P(X)$  condicional en ser parte del programa. Bernal & Peña (2011) lo definen como un promedio ponderado de las diferencias entre corchetes que utilizan la probabilidad de participación  $P(X)$  como ponderador. El estimador PSM permite predecir los resultados potenciales que no pueden ser observados por medio de una diferencia de medias de las variables de resultado entre tratamientos y controles que se encuentren en el soporte común.

En la literatura existen diferentes algoritmos de emparejamiento que se pueden implementar al PSM. La elección es un tema importante que ha sido controversial a lo largo de los años y que aún se encuentra abierto al debate; cada método posee tanto ventajas como desventajas. La discusión se centra en el *trade-off* que se genera entre varianza y sesgo dependiendo de la decisión que tome el investigador (Heckman, Ichimura & Todd, 1997). La tabla 2 resume la situación.

**Tabla 2. Algoritmos de Emparejamiento: *Trade-off* entre sesgo y varianza**

<b>Emparejamiento: vecino más cercano</b>	<b>Sesgo</b>	<b>Varianza</b>
Vecinos múltiples/ vecino 1 a 1	(+) / (-)	(-) / (+)
Con <i>caliper</i> / Sin <i>caliper</i>	(-) / (+)	(+) / (-)

Aumento (+), Disminución (-)

**Fuente:** Caliendo & Kopeinig (2005). Elaboración propia

Para la investigación, se decidió utilizar el algoritmo de vecino más cercano 1:1, sin reemplazamiento y con *caliper* (distancia máxima entre los *Propensity Score*) de 0,25 desviaciones estándar. Autores como Rubin (1973) y Stuart (2010) plantean que es uno de los métodos más comunes por su buen rendimiento y sencillez.

Una desventaja que se le adjudica a la coincidencia por vecino más cercano 1:1 es que puede descartar una gran cantidad de observaciones, generando una disminución en la precisión de los emparejamientos. Sin embargo, Cohen (1988) argumentó que la reducción suele ser mínima, debido a que la precisión se debe en mayor proporción al tamaño del grupo más pequeño y en este tipo de casos, es el grupo de control el que disminuye en tamaño buscando las coincidencias. Las coincidencias con reemplazos generarían que los controles coincidentes no sean independientes, generando que el ATT se base en un pequeño número de individuos (Stuart, 2010).

Stuart (2010) recomienda que, al tener aproximadamente tres veces más individuos en el grupo de control respecto al grupo de tratamiento, el emparejamiento de vecino más cercano sin reemplazo y a un *caliper* de 0,25 desviaciones estándar ofrece un buen rendimiento en la estimación del ATT y permite una correcta inferencia de los resultados. Resulta importante destacar que la estructura de los datos que se utilizan en la investigación cumple a cabalidad con estas características y se expondrán en mayor detalle en la siguiente sección.

De acuerdo con Khandker et al. (2010) los pasos que resumen la estimación por PSM son:

- a) Estimación de la probabilidad de participación en el programa usando las muestras de tratamiento y control por medio de un modelo *logit* o *probit*.
- b) Definición en la muestra de la región de soporte común y revisión de la propiedad de balance.
- c) Selección de mecanismo de emparejamiento entre participantes y no participantes

Bernal & Peña (2011) proponen tres pasos adicionales a realizar:

- d) Revisar que las variables observables entre los grupos de tratamiento y control sean parecidas o estén balanceadas por grupos de probabilidad predicha.
- e) Se calculan los impactos del programa como el promedio apropiadamente ponderado de la diferencia entre la variable de resultado de los tratados y los no tratados.

- f) Se calculan los errores estándar e intervalos de confianza para determinar si los impactos son estadísticamente significativos.

## 5.2 *Propensity Score Matching* con datos agrupados

En estudios de tipo educativos como la presente investigación, la estructura de los datos suele caracterizarse por estudiantes que se agrupan en cursos diferentes dificultando el método de emparejamiento, por tanto, el análisis de los resultados educativos proporciona un ejemplo clásico del requerimiento de estimaciones de efectos de tratamiento cuando los datos tienen una estructura de *cluster* donde el tratamiento se administra a nivel individual y las variables contextuales (a nivel de grupo) influyen en el mecanismo de asignación y los resultados (Arpino & Mealli, 2010).

Este apartado presenta la adaptación del marco general descrito en la sección 5.1 a una estructura de varios niveles. Se supone una estructura de datos de dos niveles donde hay  $N$  unidades en el primer nivel, indexadas por  $i$  ( $i = 1, 2, \dots, n_j$ ), que se encuentran anidadas en  $J$  unidades en el segundo nivel (*clusters*), indexadas por  $j$  ( $j = 1, 2, \dots, J$ ). Las variables de tratamiento previo pueden medirse tanto en el primer nivel ( $X$ ) como en el segundo nivel ( $Z$ ) y se asume la condición de valor de tratamiento unitario estable donde los resultados potenciales para cada unidad no se ven afectados por el tratamiento asignado a las otras unidades, independientemente del grupo al que pertenezcan, es decir, que no hay interferencia dentro y entre los grupos.

Los dos supuestos claves de PSM se reformulan de la siguiente forma:

$$Y(0), Y(1) \perp D | X, Z \quad (6)$$

$$0 < P(D=1 | X, Z) < 1 \quad (7)$$

Bajo estos supuestos, Arpino & Mealli (2010) plantean que el objetivo de la metodología es comparar las unidades tratadas y de control con características similares, ambas medidas a nivel individual y a nivel de *cluster*.

El estimador se reformula tal que:

$$e_{ij} = P(D_{ij} = 1 | (X_{ij}, Z_j)) \quad (8)$$



Rosenbaum & Rubin (1983) demostraron que el *Propensity Score* es una puntuación de equilibrio, Arpino & Cannas (2016) plantean que el equilibrio se conserva bajo la estructura *cluster*, es decir,  $(X, Z) \in D \mid e(X, Z)$ , lo que significa que en cada valor del *Propensity Score* la distribución de las covariables que lo definen es la misma en los grupos de tratamiento y control.

Una vez estimado el *Propensity Score*, se recurre a un proceso de emparejamiento por dicha probabilidad que permitirá finalmente, hallar el ATT. Arpino & Cannas (2016) resumen la estimación en las siguientes expresiones:

$$A_{rj} = \{kj' \in I_0 : \hat{e}_{kj'} = \min_{kj' \in I_0} |\hat{e}_{rj} - \hat{e}_{kj'}| < 0,25\hat{\sigma}_e\} \quad (9)$$

$$M = \{rj : A_{rj} \neq \emptyset\} \cup \left\{ \bigcup_{rj} A_{rj} \right\} \quad (10)$$

$$\widehat{ATT} = \frac{1}{card(M)} \left\{ \sum_{rj \in I_1 \cap M} \left( Y_{rj} - \sum_{kj' \in I_0} Y_{kj'} w(rj, kj') \right) \right\} \quad (11)$$

La expresión (9) muestra el conjunto de unidades emparejadas; en ella se busca el emparejamiento por distancia mínima en la probabilidad de participación entre tratamientos que pertenecen al *cluster* j y controles que pertenecen al *cluster* j que no sean mayores a 0,25 desviaciones estándar (*caliper* escogido para la investigación). La expresión (10) denota el conjunto de datos que lograron emparejarse con éxito (M), incluye tanto unidades de tratamiento como de control y permite desechar aquellas unidades que no lograron encontrar un par bajo las restricciones impuestas en (9). Finalmente, la construcción del ATT en (11) muestra un término  $w(rj, kj')$  reflejando la ponderación que se le asigne a los individuos del grupo de control en la estimación de la salida potencial no observada  $Y(0)$  para el grupo de tratados  $rj$ ; en términos generales, se traduce como la media (ponderada) de Y en unidades tratadas emparejadas menos la media (ponderada) de Y en controles emparejados.

En otras palabras, la expresión final permite observar que, para cada unidad tratada en la muestra, los algoritmos buscan la unidad de control más cercana en términos de puntaje de propensión. Si no hay una unidad de control disponible en el rango definido por el *caliper* (0,25 desviaciones estándar), la unidad tratada se descarta de la muestra de trabajo hasta lograr calcular el promedio ponderado de los ATT dentro del *cluster*, con ponderaciones

dadas por los tamaños de *cluster* en el conjunto de datos coincidentes.

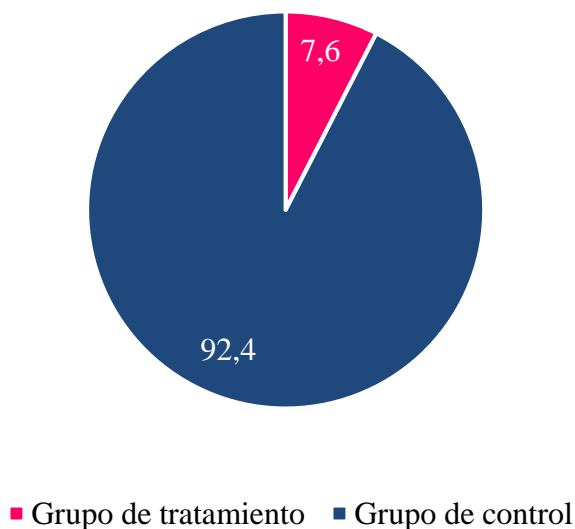
### 5.3 Estructura de los datos

Para compactar la información necesaria de la investigación, se recopilaron bases de datos de diferentes estamentos de la Universidad del Valle. Los datos de la población intervenida los proporcionó la oficina de ASES donde cuentan con los registros académicos y socioeconómicos de los estudiantes participantes del programa, especialmente para las necesidades de la investigación, los registros de los estudiantes Ser Pilo Paga II.

Por otro lado, los registros académicos del grupo de control los proporcionó la Oficina de Registro Académico y los registros socioeconómicos de esta población los proporcionó la Oficina de Planeación de la Universidad.

Se trabajó con estudiantes que tuvieran a disposición la totalidad de la información necesaria para los fines de la investigación, es decir, estudiantes que diligenciaron la encuesta socioeconómica al ingresar a la Universidad del Valle y que tuvieran reporte de calificaciones para los cuatro semestres a analizar. El grupo de tratamiento se conformó con 98 estudiantes mientras que 1.102 estudiantes constituyeron el grupo de control (Gráfico 3).

**Gráfico 3. Distribución (%) estudiantes bajo estudio no experimental**



**Fuente:** Elaboración propia

El grupo de control que se conformó para la investigación corresponde a los estudiantes regulares admitidos en la Universidad del Valle en el semestre 2016-I que no hicieron parte del programa ASES. Se podría considerar como un limitante del ejercicio el no contar con controles que obedezcan el mismo sentido en obtención de subsidios que caracteriza a los tratamientos y esto puede significar algún sesgo en la estimación; sin embargo, el ejercicio que se planteó controla las características con respecto a los atributos a nivel individual y de entorno permitiendo que los resultados arrojen evidencia a favor del planteamiento inicial como se expondrá más adelante en el desarrollo de la investigación.

La base de datos conformada es de tipo corte transversal; cuenta con características socioeconómicas y académicas para 1.200 registros de estudiantes. Las variables de carácter académico capturan el rendimiento académico de cada estudiante por cada semestre del ciclo básico (cuatro semestres). La Tabla 3 expone las variables utilizadas en la investigación mientras que sus estadísticas descriptivas se adjuntan en el Anexo 1.

**Tabla 3. Variables de análisis**

<b>VARIABLE</b>	<b>DESCRIPCIÓN DE LA VARIABLE</b>
<i>Determinantes de la deserción</i>	
<b>Esfuerzo académico</b>	Indicador que le asigna mayor ponderación al componente de asignaturas propias del plan de estudios del programa académico al que pertenece el estudiante que a las asignaturas de electivas complementarias. Además, tiene en cuenta la carga académica de los estudiantes
<b>ASES</b>	Variable dummy que toma el valor de 1 si el individuo pertenece al programa y 0 en caso contrario
<b>Nivel educativo madre</b>	Variable continua que capta el nivel de escolaridad de la madre aumentando a mayor nivel educativo alcanzado
<b>Elección de Univalle</b>	Variable categórica que indica la motivación del estudiante para escoger la Universidad del Valle como su alma máter

<i>Caracterización del individuo</i>	
<b>Sexo</b>	Variable dummy que toma el valor de 1 si el estudiante es hombre y 0 si es mujer
<b>Raza</b>	Variable categórica que indica la raza en la que se autorreconocen los estudiantes
<b>Estado civil</b>	Variable categórica que indica el estado civil del estudiante
<b>Número de hijos</b>	Variable continua, representa el número de hijos del estudiante
<b>Edad</b>	Variable continua, representa el número de años del estudiante
<b>Edad<sup>2</sup></b>	Variable continua que permite identificar rendimientos decrecientes en la variable Edad
<b>Régimen subsidiado</b>	Variable dummy que toma el valor de 1 si el estudiante hace parte del régimen subsidiado, 0 en caso contrario
<b>Acceso servicios</b>	Variable dummy que toma el valor de 1 si la vivienda del estudiante tiene acceso a todos los servicios públicos, 0 en caso contrario
<b>Muda de hogar</b>	Variable dummy que toma el valor de 1 si el estudiante proviene de otro municipio y se instala en Cali, 0 en caso contrario
<b>Programa académico</b>	Variable categórica que describe el programa académico al que pertenece el estudiante
<b>Fuente:</b> Elaboración propia	

### 5.3.1 Indicador de Esfuerzo Académico

La motivación de crear un indicador que capte el esfuerzo académico de los estudiantes es lograr refinar la forma de medición del rendimiento académico del sistema actual. Lo innovador del indicador es que permite discernir entre el esfuerzo que exigen las asignaturas propias del *pénsum* de cada programa académico y el exigido por parte de las electivas complementarias que se ofrecen en la Universidad.

El promedio ponderado con el que se evalúa el desempeño de los estudiantes en el sistema actual, tiene en cuenta el número de horas de trabajo estimadas para cada asignatura expresadas en un número de créditos. Sin embargo, las electivas complementarias llegan a alcanzar o incluso superar los créditos asignados a materias del núcleo básico y profesional del estudiante desviándose del criterio esperado. Además, bajo esta visión, se excluye el factor de carga académica definido como el número de asignaturas que matricula el estudiante.

De acuerdo con las necesidades expuestas anteriormente, se creó el indicador semestral de esfuerzo académico utilizado por Escobar et al. (2006) de tal forma que:

$$IEA_j = 2 \left( Nap * \frac{\sum P_j C_j}{\sum C_j} \right) + \left( Nec * \frac{\sum P_j C_j}{\sum C_j} \right)$$

$$IEA_j = 2(Nap * Ppp) + (Nec * Ppe)$$

Donde

*Nap*= Número asignaturas propias del pénsum

*Ppp*= Promedio ponderado propio del pénsum

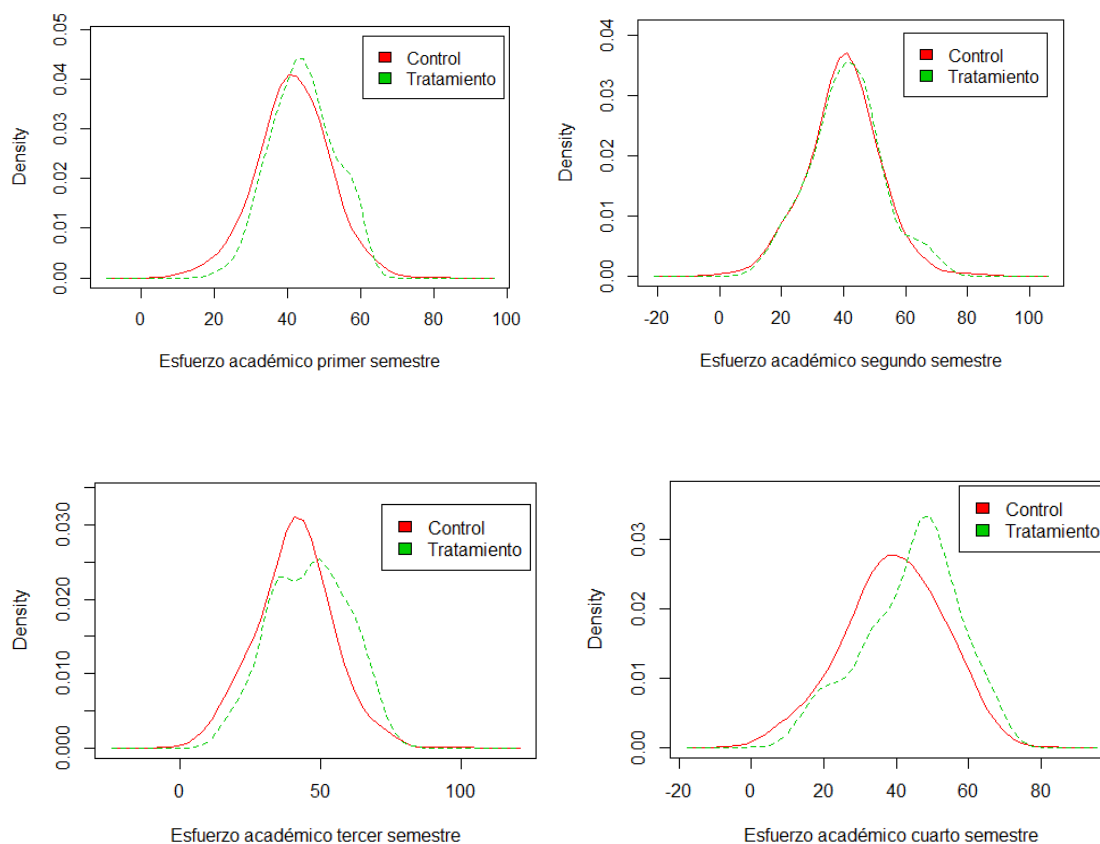
*Nec*= Número electivas complementarias

*Ppe*=Promedio ponderado electivas complementarias

En *Nap* se tienen en cuenta asignaturas básicas, profesionales y electivas profesionales que guardan relación con un mayor nivel de esfuerzo por parte del estudiante; en este sentido, se asigna una mayor ponderación al componente *pénsum*. En *Nec* se agrupan aquellas electivas complementarias que se ofrecen en diferentes facultades de la universidad, deporte formativo y constitución política de Colombia. Finalmente, los promedios ponderados contemplan los casos en que estudiantes habiliten asignaturas y se les da tratamiento como indica el Acuerdo 009 de la Universidad del Valle.

El Gráfico 4 muestra la distribución de densidad de Kernel para el indicador de esfuerzo académico construido para cada semestre académico por grupos de tratamiento y control. Se evidencia que, en los dos primeros semestres, ambos grupos guardan relación en sus distribuciones de densidad; por su parte, la distribución de los dos últimos semestres muestra que el grupo de tratamiento obtiene indicadores de esfuerzo más altos trasladándose la densidad al extremo derecho.

**Gráfico 4. Indicador de esfuerzo académico: Distribuciones de densidad semestrales**



## 5.4 Modelo

La investigación utiliza la población de estudiantes Ser Pilo Paga II que ingresaron en el primer semestre del año 2016 a la Universidad del Valle; recibiendo, por tanto, la intervención ASES. Lo anterior con el fin de analizar si el programa ASES tuvo impacto en el rendimiento académico de dicha cohorte a lo largo del ciclo básico donde se concentra la deserción temprana.

Para ello, se realiza una estimación de la probabilidad de participación de los estudiantes en el programa ASES por medio de un modelo probabilístico controlando por características socioeconómicas y académicas, dicha estimación no es más que los *Propensity Score*; posteriormente, se implementan tres modelos que contemplan diferentes métodos de coincidencias que permitirán construir el contrafactual y, al realizar una comparación entre los modelos, se podrá determinar cuál arroja el mejor resultado para la investigación.

Se consideran los siguientes modelos de coincidencia:

- a) **Emparejamiento Ingenuo (PSM-I):** modelo que empareja individuos del grupo de tratamiento y de control independientemente de la estructura de sus datos (pertenencia a un *cluster*), es decir, se ignora la estructura multinivel en la implementación de la coincidencia.
- b) **Emparejamiento por Coincidencia Pura (PSM-C):** modelo que empareja individuos del grupo de tratamiento y control respetando la estructura de los datos, coincide las unidades con los *clusters*. La coincidencia dentro de cada *cluster* garantiza automáticamente que todas las variables de nivel de *cluster* (medidas y no medidas) estén perfectamente equilibradas.
- c) **Emparejamiento Preferencial (PSM- P):** modelo que empareja individuos del grupo de tratamiento y control respetando la estructura de los datos, empieza buscando coincidencias dentro de cada *cluster*, si no tiene éxito en su respectivo grupo, las unidades de control se buscan en otros *clusters* de acuerdo con la condición (9) del apartado anterior.

#### 5.4.1 Estimación de la probabilidad de participación

El punto de partida de la metodología consiste en hallar las probabilidades de participación en el programa objeto de medición. Dichas probabilidades tienen un papel protagónico en el procedimiento ya que recogen los efectos del conjunto de covariables especificadas en la estimación y son utilizadas para emparejar a los individuos.

Se estimó la probabilidad de participación en el programa ASES por medio de un modelo probabilístico utilizando como variable dependiente la variable binaria ASES y siete variables explicativas como edad, edad al cuadrado, régimen subsidiado, acceso a servicios públicos, nivel educativo de la madre, mudanza de hogar y estrato. Cabe resaltar que como exponen Bernal & Peña (2011), solo se incluyeron en el modelo aquellas variables que alcanzaron significancia a la hora de explicar la participación en el programa a un 95% de confianza ya que la inclusión de variables que en realidad no estén asociadas con el resultado puede producir aumentos de varianza. Los resultados de la estimación se exponen en la Tabla 4.

**Tabla 4. Estimación modelos probabilísticos**

<b>Coefficientes</b>	<b>Probit</b>	<b>Logit</b>
(Intercepto)	-74,94763** (28,92365)	-158,55774** (58,81971)
muda_hogar	0,69946*** (0,14179)	1,24129*** (0,25576)
acceso_servicios	-0,41288** (0,14788)	-0,74579** (0,26985)
edad	8,02558** (2,97749)	16,95719** (6,07446)
edad2	-0,21557** (0,07655)	-0,45440** (0,15671)
niveleducativo_madre	-0,06950* (0,03125)	-0,12084* (0,05904)
estrato	-0,19009* (0,07926)	-0,34734* (0,15159)
regimen_subsidiado	0,34577* (0,14984)	0,62883* (0,26985)
AIC	496,2973	498,4409
BIC	537,0179	539,1615

Errores estándar en paréntesis

\*p<0,1, \*\*p<0,5, \*\*\*p<0,01

**Fuente:** Elaboración propia

Como se planteó en Bernal & Peña (2011), la elección del modelo de probabilidad no es problemático en este tipo de metodología ya que tanto el modelo *logit* como el *probit* arrojan resultados similares como se logra observar en la Tabla 4. No obstante, se utilizaron los criterios de información Akaike (AIC) y Bayesiano (BIC) para determinar el mejor modelo en términos de ajuste para estimar la probabilidad de participación en el programa; el modelo *probit* tiene los menores valores en ambos criterios y es el elegido.

Se encuentra que variables como sexo y raza no son determinantes de la participación en ASES y no se tienen en cuenta en la estimación; esto muestra que el diseño del programa está abierto para cualquier estudiante independientemente de su sexo y su color de piel. Por el contrario, se deja en evidencia que el programa está focalizado para aquellos estudiantes que cuentan con factores de riesgo de deserción tales como pertenecer a un estrato socioeconómico bajo, tener carencias en el acceso total a servicios públicos, pertenecer al régimen subsidiado, moverse de su ciudad de origen para instalarse en una nueva ciudad como Cali y un nivel educativo de la madre bajo. Respecto a la edad, se evidencian rendimientos decrecientes por medio del cambio de signo de la variable edad2; a medida que aumente la edad hay una mayor probabilidad de participar en el programa, pero hay un punto en que tener mayor edad disminuye la probabilidad de participación guardando relación con lo visto en las estadísticas descriptivas de esta variable.



### 5.4.2 Resultados modelos *Propensity Score Matching*

Dos de los modelos que se estiman a partir de los tres métodos de emparejamiento, incluyen siete variables de nivel individual y una variable de nivel de *cluster* (programa académico) mientras que un modelo incluye solamente las siete variables de nivel individual. En este sentido, se realiza la comparación de tres modelos alternativos para la estimación de PSM que permite equilibrar estas variables entre los grupos de tratamiento y control; para cada unidad tratada en la muestra, los algoritmos buscan la unidad de control más cercana en términos de puntaje de propensión, en caso de no haber una unidad de control disponible en el rango definido por el *caliper*, la unidad tratada se descarta de la muestra de trabajo ya que se encuentra por fuera del soporte común.

En la estimación se utilizó como variable de resultado el indicador de esfuerzo académico individualmente para cada semestre, ASES como variable que indica los tratamientos y controles, programa académico como un vector que describe la estructura de *clusters* y se estructuró el comando de forma que se calculara los ATT con emparejamiento uno a uno, *caliper* de 0,25 desviaciones estándar, sin reemplazamiento y se designó que en caso de que haya coincidencias en los *propensity score* para varios estudiantes, la elección se haga de forma aleatoria. Los resultados se exponen en la Tabla 5.

**Tabla 5. Modelos**

		PSM-I	PSM-C	PSM-P
<b>Primer semestre</b>	<b>Num unidades emparejadas</b>	91	70	98
	<b>Num unidades no emparejadas</b>	7	28	0
	<b>% muestra emparejada</b>	92,9%	71,4%	100%
	<b>ATT</b>	0,8781209	0,2913143	0,5569082
	<b>Error estándar</b>	1,172493	1,027734	1,057286
<b>Segundo semestre</b>	<b>Num unidades emparejadas</b>	91	70	98
	<b>Num unidades no emparejadas</b>	7	28	0
	<b>% muestra emparejada</b>	92,9%	71,4%	100%
	<b>ATT</b>	0,4171758	-0,2813286	0,2505714
	<b>Error estándar</b>	1,72815	1,225108	1
<b>Tercer semestre</b>	<b>Num unidades emparejadas</b>	91	70	98
	<b>Num unidades no emparejadas</b>	7	28	0
	<b>% muestra emparejada</b>	92,9%	71,4%	100%
	<b>ATT</b>	4,026352	3,601029	3,859337
	<b>Error estándar</b>	1,815169	1,424436	1,418441

<b>Cuarto semestre</b>	<b>Num unidades emparejadas</b>	91	70	98
	<b>Num unidades no emparejadas</b>	7	28	0
	<b>% muestra emparejada</b>	92,9%	71,4%	100%
	<b>ATT</b>	4,54111	2,479443	3,301673
	<b>Error estándar</b>	1,713359	1,114198	1,634543

**Fuente:** Elaboración propia

Se encontró un patrón en los resultados de cada alternativa de modelo propuesto para la investigación. El programa ASES tuvo un impacto positivo en el rendimiento académico de los estudiantes Ser Pilo Paga en el segundo año de intervención.

Ahora bien, al contrastar cada modelo y los resultados que arrojó, se encontró que el modelo ingenuo (PSM-I) corresponde a una visión que no logra capturar el objetivo de la investigación debido a que utiliza un algoritmo que empareja estudiantes independientemente del programa académico al que pertenezcan; el resultado dejó en evidencia que se sobreestima el efecto del programa cuando se emparejan individuos que no son estrictamente comparables.

El modelo preferencial (PSM-P) arroja mejores resultados en función del objetivo de la investigación, considerando que se captura la estructura de los datos al involucrar en el algoritmo la variable de *cluster*. Se encuentra que el efecto del programa ASES en el rendimiento académico de los estudiantes intervenidos es positivo durante el segundo año de estudio, no obstante, puede seguir sobreestimándose tal efecto al mezclar estudiantes de diferentes programas académicos cuando no se halle coincidencia en el *cluster* y, además, se encuentra una mayor complejidad para realizar una interpretación de los resultados.

Por último, el modelo de coincidencia pura (PSM-C) posee varios atributos, entre ellos, se encuentra que no solo involucra la estructura de *cluster* de los datos, sino que el algoritmo respeta la característica de agrupación generando emparejamientos que se pueden asegurar que son comparables.

El modelo escogido para la investigación es el modelo de coincidencia pura (PSM-C) y la decisión se justifica en el *trade-off* entre sesgo y varianza. Al querer encontrar el efecto del programa ASES en el rendimiento académico de los estudiantes, se buscaron los individuos que guardaran la más estrecha semejanza para construir el contrafactual deseado; es por ello, que se decidió realizar un emparejamiento por *cluster*, sin reemplazamiento, 1 a 1, que al final permitiera analizar e interpretar los resultados de

forma clara donde se tuviera certeza que son individuos comparables guiados por las pruebas de sensibilidad de Rosenbaum, balanceo de emparejamiento, p-valor de Kolmogorov-Smirnov y t-student (Anexo 2)

Los resultados dejan en evidencia que el impacto de ASES en el esfuerzo académico de los estudiantes de la cohorte 2016-I para el tercer semestre fue de 3,60 puntos por encima de los estudiantes sin intervención y, por último, en el cuarto semestre ASES mejoró el rendimiento de los estudiantes intervenidos en 2,47 puntos.

### **5.4.3 Prueba de sensibilidad de Rosenbaum**

Rosenbaum (2002) contempló que solo los experimentos aleatorios garantizan estar libres de sesgos de selección y, por el contrario, que estudios no experimentales como el PSM recurren a utilizar entre sus supuestos, la independencia condicional que se traduce en asumir que no existen factores no observados que determinen la participación en el programa y a su vez la variable de resultado potencial del modelo; de esta forma, las covariables observables utilizadas pueden explicar la participación en el programa. Con base en lo anterior, desarrolló pruebas de sensibilidad para las estimaciones no aleatorizadas realizadas por PSM.

Las pruebas de sensibilidad de Rosenbaum buscan analizar si se logró coincidir correctamente los datos como si se hubiera asignado aleatoriamente el tratamiento de tal forma que no se hallen diferencias entre los grupos de tratamiento y control siendo constante la probabilidad de recibir el tratamiento. Para ello, utiliza un parámetro de sensibilidad  $\Gamma$  que mide el grado de desviación de la asignación aleatoria de tratamiento; en otras palabras, dos sujetos que cuenten con iguales características observadas pueden diferir en su probabilidad de recibir el tratamiento en un factor de  $\Gamma$  que para efectos de la prueba, se pondrá a variar mediante el uso de *odds-ratio*.

Keele (2010) se cuestiona respecto a qué tan grandes deberían ser las diferencias en el parámetro  $\Gamma$  para cambiar nuestra inferencia básica teniendo en cuenta que, el tamaño de  $\Gamma$  determina las diferencias en las probabilidades de recibir el tratamiento. Plantea que, en un experimento aleatorio, la aleatorización del tratamiento asegura la no existencia de sesgo no observado ( $\Gamma=1$ ) y, en las ciencias sociales, dichos sesgos no suelen ser robustos por lo que se utilizaría en la prueba variaciones en el factor de 1 a 2.

Rosenbaum (2002) expone entre sus pruebas de sensibilidad la estimación del punto de Hodges-Lehmann en pares emparejados; esta puede ser interpretada como una aproximación a diferencias entre las medianas de los grupos de tratamiento y control (Keele, 2010). Siguiendo las recomendaciones de Keele (2010), se estableció como valor máximo del factor  $\Gamma = 1,5$  con incrementos o variaciones de 0,1 que son los parámetros regularmente utilizados en estudios de ciencias sociales y los resultados de la prueba se exponen en la Tabla 6.

**Tabla 6. Prueba de sensibilidad de Rosenbaum**

<b>Prueba de sensibilidad de Rosenbaum</b>			
<b>Estimación del punto Hodges-Lehmann</b>			
	<b>Gamma(<math>\Gamma</math>)</b>	<b>Límite inferior</b>	<b>Límite superior</b>
<b>Primer semestre</b>	<b>1</b>	1,5202e-05	1,5202e-05
	<b>1,1</b>	-0,3000	0,3000
	<b>1,2</b>	-0,6000	0,6000
	<b>1,3</b>	-0,8000	0,9000
	<b>1,4</b>	-1,1	1,2
	<b>1,5</b>	-1,3	1,5
<b>Segundo semestre</b>	<b>1</b>	0,2960	0,2960
	<b>1,1</b>	-0,2040	0,7960
	<b>1,2</b>	-0,6040	1,096
	<b>1,3</b>	-1,004	1,496
	<b>1,4</b>	-1,404	1,796
	<b>1,5</b>	-1,804	1,996
<b>Tercer semestre</b>	<b>1</b>	3,9725	3,9725
	<b>1,1</b>	3,3725	4,5725
	<b>1,2</b>	2,8725	5,0725
	<b>1,3</b>	2,3725	2,5725
	<b>1,4</b>	2,0725	2,8725
	<b>1,9</b>	0,0725	7,5725
	<b>2</b>	-0,3274	7,9725
<b>Cuarto semestre</b>	<b>1</b>	3,1976	3,1976
	<b>1,1</b>	2,4976	3,9976
	<b>1,2</b>	1,7976	4,3976
	<b>1,3</b>	1,0976	4,9976
	<b>1,4</b>	0,4976	5,5976
	<b>1,5</b>	-0,0024	6,0976

**Fuente:** Elaboración propia

De acuerdo con la información arrojada por la prueba para las diferentes variables de resultado, se encontró que el sesgo de selección no observado varía significativamente entre los diferentes resultados.

Por un lado, según la prueba de límites de Hodges-Lehmann el efecto promedio del tratamiento sobre el rendimiento académico de los estudiantes en el primer y segundo semestre es sensible ante variables no observables que contemplan sesgos no observados; características no observadas que aumenten los *odds-ratio* en al menos 10% podría sesgar el impacto estimado para estos semestres que resultaron no tener un efecto significativo. A diferencia de los primeros semestres, el impacto estimado en el tercer y cuarto semestre es menos sensible ante variaciones en factores no observables; muestra de ello, el análisis de sensibilidad de Hodges-Lehmann muestra que los *odds-ratio* tendrían que aumentar hasta 50% para sesgar el resultado del cuarto semestre y hasta el 100% (duplicarse) para el tercer semestre.

Tal como lo menciona Rosenbaum & Krieger (1990), Aakvik (2001), Watson (2005), Becker & Caliendo (2007) y Clément (2011), las pruebas de sensibilidad llevan la selección no observada a los extremos o peores escenarios para analizar cómo los sesgos podrían alterar las estimaciones; sin embargo, el hecho de encontrar que un resultado es sensible a pequeños sesgos no es afirmar que realmente existan ni que se asegure evidencia contra el efecto del tratamiento que se ha hallado.

## 6. CONCLUSIONES

Los resultados de la investigación sugieren que el rendimiento académico de los estudiantes intervenidos por ASES fue superior respecto al rendimiento académico de los estudiantes no intervenidos. En los primeros dos semestres que se pueden considerar como la etapa de transición y adecuación a las actividades y herramientas que brinda el programa, no se encontró significancia estadística. En cambio, a partir del primer año de intervención, los resultados del acompañamiento socioeducativo que instauró ASES en los estudiantes Ser Pilo Paga muestran grandes cambios positivos.

Se encontró que variables como sexo, raza, estado civil y número de hijos no son determinantes de la participación en el programa ASES. Por su parte, variables como estrato socioeconómico, acceso a servicios públicos, nivel educativo de la madre, traslado de lugar de origen y la edad sí son factores que influyen en la probabilidad de participación.

En términos metodológicos, la estimación por *Propensity Score Matching* mostró ser flexible al permitir ser aprovechada para estudios de economía de la educación donde las distribuciones de los datos guardan estructuras de *cluster* como los son los programas académicos, cursos de colegios e incluso instituciones educativas como colegios o universidades, generando resultados que fácilmente pueden ser interpretados.

Como observación metodológica se podría plantear que es posible que los resultados obtenidos puedan estar sesgados al no tener en el grupo de control una característica que se asemeje al subsidio que reciben los tratamientos. No obstante, los resultados que se encuentran con respecto al rendimiento académico en el primer año, no logran diferenciarse significativamente entre los grupos; los efectos del programa se ven claramente identificados en el segundo año de intervención.

Un ejercicio más refinado podría buscar el equivalente al subsidio en los controles para analizar si los resultados sufren alteraciones. Sin embargo, si lo que se pretende es analizar el rendimiento académico de los estudiantes, los hallazgos en el primer año resultan ser contraevidentes en busca de esa solución. El efecto del subsidio se diluye al comprobar que el desempeño mejora en el tiempo.

Para futuros estudios, resulta pertinente involucrar al análisis o evaluar específicamente el impacto de la herramienta GeoCALÍzate que instauró ASES para geolocalizar el lugar de residencia de los estudiantes intervenidos y ayudar a ubicar una vivienda con las mejores condiciones en términos de costo, rutas de acceso, tiempo de desplazamiento y cercanía a la Universidad para estudiantes que provinieran de un lugar de origen diferente a la ciudad de Cali. En la investigación no fue posible involucrar la variable espacial de la distancia del hogar a la universidad ya que solo se dispone de estos datos para los estudiantes ASES y no para el resto de la comunidad universitaria.

Se esclarece el hecho de que la estrategia ASES es un mecanismo efectivo para el acercamiento de los estudiantes al éxito académico, evidenciado en este ejercicio los efectos positivos en el ciclo básico de las distintas carreras. Recordemos que mejorar el

desempeño académico en los cuatro primeros semestres resulta vital pues los indicadores de deserción se concentran en esta etapa. Con base en ello se recomienda la aplicación del Programa ASES como un lineamiento de política universitaria para mejorar el desempeño académico de los estudiantes.

En este sentido, ante los resultados observados en el desempeño académico, los estudiantes beneficiarios del Programa Ser Pilo Paga podrán aumentar la probabilidad de éxito académico y, por esta vía, lograr la condonación pretendida del crédito otorgado por el Ministerio de Educación Nacional. Se reitera que estos resultados apuntan a que la Universidad debería contemplar la toma de decisiones orientadas a ampliar la cobertura del programa ASES. El escalamiento del Programa podría recurrir a criterios de focalización usando variables como las halladas en este ejercicio. Si los resultados de éxito académico persisten, se podría demostrar, con evidencias claras, los beneficios manifiestos en la reducción del fracaso académico mejorando así los indicadores de permanencia, reduciendo el rezago y mejorando las tasas de graduación.

## 7. REFERENCIAS

- Aakvik, A. (2001). Bounding a matching estimator: the case of a Norwegian training program. *Oxford bulletin of economics and statistics*, 63(1), 115-143.
- Ajzen, I. & Fishbein, M. (1975). Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research. *Reading MA AddisonWesley*, (August), 480.
- Angrist, J., Bettinger, E., & Kremer, M. (2006). Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia. *American Economic Review*.
- Arpino, B., & Cannas, M. (2016). Propensity score matching with clustered data. An application to the estimation of the impact of caesarean section on the Apgar score. *Statistics in medicine*, 35(12), 2074-2091.
- Arpino, B., & Mealli, F. (2010). The specification of the propensity score in multilevel observational studies. *Computational Statistics & Data Analysis*, 55(4), 1770-1780.

- Attinasi, L. C. (1986). Getting in: Mexican American Students' perceptions of their college-going behavior with implications for their freshman year persistence in the University. In *ASHE, 1986 Annual Meeting Paper. San Antonio, TX, EE. UU. (ERIC N.º 268 869)*.
- Barbosa, F. A., Castillo, N. S., & Vásquez, M. L. (2018). Estrategia de Acompañamiento y Seguimiento Estudiantil en la Universidad del Valle (Colombia).
- Barrera-Osorio, F. (2007). The Impact of Private Provision of Public Education: Empirical Evidence from Bogotá's Concession Schools. *Policy Research Working Paper Series*, 1–30.
- Becker, S. O., & Caliendo, M. (2007). Mhbounds-sensitivity analysis for average treatment effects.
- Bean, J. P. (1985). Interaction Effects Based on Class Level in an Explanatory Model of College Student Dropout Syndrome. *American Educational Research Journal*, 22(1), 35–64.
- Bernal, R., & Peña, X. (2011). *Guía práctica para la evaluación de impacto*. Ediciones Uniandes-Universidad de los Andes.
- Braxton, J. M., Milem, J. F., & Sullivan, A. S. (2000). The influence of active learning on the college student departure process: Toward a revision of Tinto's theory. *The Journal of Higher Education*, 71(5), 569–590.
- Caliendo, M., & Kopeinig, S. (2005). Some practical guidance for the implementation of Propensity Score Matching. Bonn, Germany. *Institute for the Study of Labor*, Discussion Paper,(1588).
- Clément, M. (2011). Remittances and household expenditure patterns in Tajikistan: A propensity score matching analysis.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Lawrence Earlbaum Associates*; Hillsdale, NJ.
- De Escobar, V., de Tirado, N. D., de Santana, O. F., & Hernández, D. (2005). Estudio sobre la deserción y repitencia en la educación superior en Panamá. Consejo de Rectores.
- Durkheim, E. (1951). Suicide, a study in sociology. *New York Free Press*, 3, 405.
- Eccles, J. S., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, Values, and Academic Behaviors. *Achievement and Achievement Motivation*.
- Escobar, J., Largo, E., & Pérez, C. A. (2006). Factores asociados a la deserción y permanencia estudiantil en la Universidad del Valle (1994-2006). *Cali, Universidad del Valle, Vicerrectoría Académica, Facultad de Ciencias Sociales y Económicas, Centro de Investigación y documentación socioeconómica-CIDSE*.



- Escobar, J., Largo, E., & Pérez, C. A. (2008). Los Diez Lineamientos: Recomendaciones de Política para Incidir en la Deserción y Permanencia de los Estudiantes de la Universidad del Valle. *Cali, Universidad del Valle, Vicerrectoría Académica, Facultad de Ciencias Sociales y Económicas, Centro de Investigación y documentación socioeconómica-CIDSE*.
- Ethington, C. A. (1990). A psychological model of student persistence. *Research in Higher Education*, 31 (3), pp. 266-269.
- Girón, L., & González, D. (2005). Determinantes del rendimiento académico y la deserción estudiantil, en el programa de economía de la Pontificia Universidad Javeriana de Cali. *Economía, Gestión Y Desarrollo*, 3, 173–201.
- Guzmán, C., Durán, D., Franco, J., Castaño, E., Gallón, S., Gómez, K., & Vásquez, J. (2009). Deserción estudiantil en la educación superior colombiana. Metodología de seguimiento, diagnóstico y elementos para su prevención. *Ministerio de Educación Nacional. Colombia*.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, 64(4), 605-654.
- Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. In *Handbook of labor economics* (Vol. 3, pp. 1865-2097). Elsevier.
- Himmel, E. (2002). Modelos de análisis de la deserción estudiantil en la educación superior. *Revista calidad de la educación*, 17, 91-108.
- Joffe, M. M., & Rosenbaum, P. R. (1999). Invited commentary: propensity scores. *American Journal of Epidemiology*, 150(4), 327–33.
- John, E. P., Cabrera, A. F., Nora, A., & Asker, E. H. (1999). Economic influences on persistence reconsidered. *Reworking the student departure puzzle*, 29-47.
- Keele, L. (2010). An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. *White Paper. Columbus, OH*, 1-15.
- Khandker, S. R., Koolwal, G. B., & Samad, H. a. (2010). *Handbook on Impact Evaluation: Quantitative Methods and Practices. Learning*.
- Lassibille, G., & Navarro Gómez, M. L. (2004). Manual de economía de la educación. Madrid: Ediciones Pirámide.
- Lopera, C. (2008). Determinantes de la deserción universitaria en la Facultad de Economía Universidad del Rosario. *Borradores de Investigación*, (95), 26.
- Luellen, J. K., Shadish, W. R., & Clark, M. H. (2005). Propensity scores: an introduction and experimental test. *Evaluation Review*, 29(6), 530–58.

- Ministerio de Educación Nacional de Colombia. (2014). Informe mensual sobre el soporte técnico y avance del contrato para garantizar la alimentación, consolidación, validación y uso de la información del SPADIES. Recuperado de: [http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-254702\\_Informe\\_determinantes\\_desercion.pdf](http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-254702_Informe_determinantes_desercion.pdf)
- Morín, E. (1999). Los siete saberes necesarios para la educación del futuro. *Revista De Innovación E Investigación Educativa*, 60.
- Porto, A., & Gresia, L. Di. (2004). Rendimiento de estudiantes universitarios y sus determinantes. *Revista de Economía Y Estadística*, 42(1), 93–113.
- Ramírez, A. (2013). Evaluación de Impacto del Proyecto ICARO en la Deserción Escolar (Impact Evaluation of the Project ICARO on School Desertion).
- Rosenbaum, P. R. (2002). Observational studies. In *Observational studies* (pp. 1-17). Springer, New York, NY.
- Rosenbaum, P. R., & Krieger, A. M. (1990). Sensitivity of two-sample permutation inferences in observational studies. *Journal of the American Statistical Association*, 85(410), 493-498.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics*, 159-183.
- Spady, W. G. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, 1(1), pp. 64- 85.
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 25(1), 1.
- Tillman, S., C.A (2002) Barriers to Student Persistence in Higher Education.
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research*, 45(1), 89-125.
- Tinto, V. (1989). Definir la deserción: una cuestión de perspectiva. *Revista de educación superior*, 71(18), 1-9.
- Vásquez, J., Castaño, E., Gallón, S., & Gómez, K. (2003). Determinantes de la deserción estudiantil en la Universidad de Antioquía. *Centro de Investigaciones Económicas*, 4(1), 1–40.
- Watson, I. (2005). The earnings of casual employees: The problem of unobservables. In *HILDA Survey Research Conference. The University of Melbourne* (Vol. 29).
- Weidman, J. C. (1989). Undergraduate socialization: A conceptual approach. In *Higher education: Handbook of theory and research* (pp. 289–322).

## Anexo 1. Estadísticas descriptivas

Estadísticas descriptivas				
	Tratamientos		Controles	
	freg abs	freg rel	freg abs	freg rel
<b>Sexo</b>				
Hombre	53	54,1%	643	58,3%
Mujer	45	45,9%	459	41,7%
<b>Edad</b>				
18 años	9	9,2%	59	5,4%
19 años	60	61,2%	307	27,9%
20 años	27	27,6%	241	21,9%
21 años	1	1,0%	119	10,8%
22 años	1	1,0%	96	8,7%
Otros	0	0,0%	280	25,4%
<b>Estado civil</b>				
Soltero	98	100%	1067	96,8%
Unión libre	0	0%	26	2,4%
Casado	0	0%	7	0,6%
Separado	0	0%	1	0,1%
Viudo	0	0%	1	0,1%
<b>Número de hijos</b>				
<b>0</b>	98	100%	1071	97,2%
<b>1</b>	0	0%	25	2,3%
<b>2</b>	0	0%	3	0,3%
<b>3</b>	0	0%	1	0,1%
<b>4</b>	0	0%	1	0,1%
<b>5</b>	0	0%	1	0,1%
<b>Raza</b>				
Blanco, Mestizo	85	86,7%	937	85,0%
Negro,mulato, afrocolombiano	8	8,2%	119	10,8%
Indígena	4	4,1%	42	3,8%
Otro	1	1,0%	4	0,4%
<b>Estrato socioeconómico</b>				
Estrato 1	48	49,0%	262	23,8%
Estrato 2	31	31,6%	436	39,6%
Estrato 3	18	18,4%	302	27,4%
Estrato 4	1	1,0%	70	6,4%
Estrato 5	0	0,0%	27	2,5%
Estrato 6	0	0,0%	5	0,5%

<b>Traslado de lugar de origen</b>				
Se traslada	50	51,0%	172	15,6%
No se traslada	48	49,0%	930	84,4%
<b>Nivel educativo madre</b>				
No tuvo escuela	1	1,0%	12	1,1%
Primaria incompleta	8	8,2%	88	8,0%
Primaria completa	9	9,2%	66	6,0%
Bachillerato incompleto	17	17,3%	150	13,6%
Bachillerato completo	44	44,9%	321	29,1%
Tecnólogo/ Técnico incompleto	0	0,0%	23	2,1%
Tecnólogo/ Técnico completo	15	15,3%	175	15,9%
Universitario incompleto	0	0,0%	49	4,4%
Universitario completo	2	2,0%	162	14,7%
Posgrado	2	2,0%	56	5,1%
<b>Elección de Univalle</b>				
Su bajo costo de matrícula	6	6,1%	202	18,3%
Única universidad que ofrece el programa	2	2,0%	30	2,7%
Riqueza cultural	4	4,1%	58	5,3%
Prestigio y reconocimiento	86	87,8%	812	73,7%
<b>Acceso servicios públicos</b>				
Insuficiente	54	55,1%	260	23,6%
Completa	44	44,9%	842	76,4%

---

**Fuente:** Elaboración propia

## **Anexo 2. Prueba de balanceo de emparejamiento**

Se calculó el p-valor de *bootstrap* de la prueba de Kolmogorov-Smirnov para las variables continuas utilizando la hipótesis de que las densidades de probabilidad tanto para el grupo tratado como para el de control son las mismas. Valores mayores a 0,05 no rechazan la hipótesis nula y denotan que las densidades son iguales; por ende, están bien emparejadas. Otra de las pruebas realizadas para todas las variables del emparejamiento consistió en hallar la relación varianza tratamiento/control; entre más cercano el valor a 1(balanceo perfecto) mejor se considera el balanceo realizado. Además, se calculó la prueba de p-valor *t-student* que tiene por hipótesis nula la inexistencia de diferencias entre medias de los grupos, siendo positivo para el emparejamiento que el valor calculado no sea significativo.

***** (V1) Régimen subsidiado *****		
	Before Matching	After Matching
mean treatment	0,33673	0,3
mean control	0,18512	0,32857
std mean diff	31,918	-6,1901
mean raw eQQ diff	0,15306	0,028571
med raw eQQ diff	0	0
max raw eQQ diff	1	1
mean eCDF diff	0,075808	0,014286
med eCDF diff	0,075808	0,014286
max eCDF diff	0,15162	0,028571
var ratio (Tr/Co)	1,4945	0,9519
T-test p-value	0,0027056	0,63831
***** (V2) Edad *****		
	Before Matching	After Matching
mean treatment	19,235	19,257
mean control	21,426	19,314
std mean diff	-326,91	-7,9811
mean raw eQQ diff	2,3061	0,085714
med raw eQQ diff	1	0
max raw eQQ diff	24	1
mean eCDF diff	0,087164	0,017143
med eCDF diff	0,030853	0,014286
max eCDF diff	0,42696	0,028571
var ratio (Tr/Co)	0,037279	1,1379
T-test p-value	<2,22e-16	0,61242
KS Bootstrap p-value.	<2,22e-16	0,955
KS Naive p-value	1,12e-14	1
KS Statistic	0,12696	0,028571

***** (V3) Estrato *****		
	Before Matching	After Matching
mean treatment	1,7143	1,8857
mean control	2,255	2,0143
std mean diff	-67,632	-15,906
mean raw eQQ diff	0,54082	0,12857
med raw eQQ diff	1	0
max raw eQQ diff	2	2
mean eCDF diff	0,090118	0,025714
med eCDF diff	0,055697	0,028571
max eCDF diff	0,25205	0,042857
var ratio (Tr/Co)	0,64247	0,69378
T-test p-value	5,20E-09	0,32329
KS Bootstrap p-value.	<2,22e-16	0,904
KS Naive p-value	2,16e-05	1
KS Statistic	0,25205	0,042857

***** (V4) Edad <sup>2</sup> *****		
	Before Matching	After Matching
mean treatment	370,42	371,34
mean control	471,14	373,49
std mean diff	-385,49	-7,6568
mean raw eQQ diff	109,1	3,3714
med raw eQQ diff	39	0
max raw eQQ diff	1632	43
mean eCDF diff	0,087164	0,017143
med eCDF diff	0,030853	0,014286
max eCDF diff	0,42696	0,028571
var ratio (Tr/Co)	0,021154	1,1481
T-test p-value	<2,22e-16	0,62434
KS Bootstrap p-value.	<2,22e-16	0,955
KS Naive p-value	1,12e-14	1
KS Statistic	0,42696	0,028571

***** (V5) Acceso servicios *****		
	Before Matching	After Matching
mean treatment	0,44898	0,55714
mean control	0,76407	0,58571
std mean diff	-63,024	-5,7107
mean raw eQQ diff	0,31633	0,028571
med raw eQQ diff	0	0
max raw eQQ diff	1	1
mean eCDF diff	0,15754	0,014286
med eCDF diff	0,15754	0,014286
max eCDF diff	0,31509	0,028571
var ratio (Tr/Co)	1,3853	1,0168
T-test p-value	2,07e-08	0,65566

***** (V6) Nivel educativo madre *****		
	Before Matching	After Matching
mean treatment	3,8469	4,0714
mean control	4,775	3,9
std mean diff	-54,553	9,2568
mean raw eQQ diff	0,93878	0,6
med raw eQQ diff	1	0
max raw eQQ diff	4	2
mean eCDF diff	0,092939	0,055556
med eCDF diff	0,051715	0,014286
max eCDF diff	0,22808	0,14286
var ratio (Tr/Co)	0,54074	0,58822
T-test p-value	1,76e-06	0,56556
KS Bootstrap p-value.	<2,22e-16	0,192
KS Naive p-value	1,72e-04	0,47271
KS Statistic	0,22808	0,14286

***** (V7) Muda hogar *****		
	Before Matching	After Matching
mean treatment	0,5102	0,41429
mean control	0,15608	0,42857
std mean diff	70,477	-2,8793
mean raw eQQ diff	0,34694	0,014286

med raw eQQ diff	0	0
max raw eQQ diff	1	1
mean eCDF diff	0,17706	0,0071429
med eCDF diff	0,17706	0,0071429
max eCDF diff	0,35412	0,014286
var ratio (Tr/Co)	1,915	0,99083
T-test p-value	5,76e-10	0,82785

**Fuente:** Elaboración propia



